

Tracking and Prediction of Large-scale Organized Tropical Convection by Spectrally
Focused Two-Step Space-Time EOF Analysis

Paul E. Roundy
University at Albany

State University of New York

Submitted to The Quarterly Journal of the Royal Meteorological Society July 2011

Corresponding Author Address:

Paul E. Roundy
University at Albany
Department of Atmospheric and Environmental Sciences
DEAS-ES 351
Albany, NY 12222

Abstract

An EOF analysis in time and space is applied to extract the coherent signals of convectively coupled equatorial waves, intraseasonal oscillations, and other disturbances from unfiltered satellite outgoing longwave radiation anomaly data. The algorithm produces a basis of time indices for the coherent signals in selected bands of the zonal wave number frequency domain and also generates reduced noise versions of wave number frequency filtered data applicable in real time. Multiple linear regression is applied to forecast the time indices of each wave number frequency band, and the predicted indices are applied to reconstruct the predicted filtered OLR fields. A cross validation analysis demonstrates that the predicted MJO signals exhibit skill to 25 days across the global tropics and beyond 30 days across some of the higher latitudes of the tropics.

Key Words: Madden Julian oscillation, El Niño/Southern oscillation, convectively coupled equatorial waves, intraseasonal prediction

1. Introduction

Prediction at the sub seasonal weather-climate interface is on the frontier of investigation in atmospheric sciences. Spectral peaks in proxies for atmospheric moist deep convection in the tropics that project well above others in their neighborhoods suggest the presence of coherent intraseasonal signals that might yield predictability by empirical means. The associated signals contribute substantially to the sensible weather in the tropical atmosphere and are also associated with signals in the extratropical circulation (Wallace and Gutzler 1981; Ferranti et al. 1990; Weickmann et al 1997; Mo and Higgins 1998; Hendon et al. 2000; Higgins et al. 2000; Jones et al. 2000; Nogues-Paegle et al. 2000; Mo 2000; Branstator 2002; Jones et al. 2004a,b; Weickmann and Berry 2007). The associations between these signals in the tropical atmosphere and extratropical circulations suggest that although such pronounced spectral peaks are less apparent in the mid latitudes, the portions of the mid latitude signals that are coherent with the tropical convective modes might yield a predictable intraseasonal background state to the synoptic weather.

The 30-60 day Madden Julian oscillation (MJO, Madden and Julian 1994; Zhang 2005) acts along with interannual patterns such as the El Niño/Southern Oscillation (ENSO) to modulate the organization and evolution of such moist deep convection and the global atmospheric circulation (Roundy et al. 2010). Synoptic to planetary scale waves couple to convection and modulate its evolution on higher frequencies (Kiladis et al. 2009). These waves include convectively coupled equatorial Rossby (ER), Kelvin, mixed Rossby-gravity (MRG), and easterly waves (e.g., Kiladis et al. 2005). All of these

disturbances evolve together as part of the same nonlinear system. Changes in one mode influence the background conditions felt by the others. Nevertheless, some of these wave and climate modes evolve quasi-systematically on intraseasonal timescales (herein considered roughly 10 to 90 days) with linear signals apparently dominating their evolution. The term “linear” herein implies that if a particular convective mode can be described by a basis of time series, then a future state of that mode is approximately a linear combination of the present signals in that basis. A portion of the nonlinear signal might also yield predictability. For example, nonlinear interaction with the seasonally evolving base state is associated with changes in the structure and propagation characteristics of the MJO and convectively coupled waves. When such modulation occurs in similar ways across many years, statistical models can be trained to diagnose and predict its effects.

Recently, Roundy and Schreck (2009, hereafter RS09) developed a statistical system for identifying and tracking the signals of convective modes proximate to selected pronounced spectral peaks. They calculated the leading time extended empirical orthogonal functions (EEOFs) of outgoing longwave radiation (OLR) anomalies filtered in time and space to emphasize the signals of selected modes of organized convection. They found that by projecting unfiltered OLR anomalies onto the EEOF patterns of the filtered anomalies they could extract signals associated with the target modes. The time extensions include only the recent past (i.e., no future data are required) so that the algorithm can be applied in real time. The EEOF patterns provide general synoptic-dynamic models of the progression of each of the target modes because they include zonal, meridional, and temporal degrees of freedom.

Other authors have applied EOF methods to track the MJO in real time. For example, Wheeler and Hendon (2004, hereafter WH04) combined OLR and zonal wind data averaged over 15°N to 15°S in an EOF analysis, and they labeled the resulting PC time series the real time multivariate MJO (RMM) indices. By averaging over latitude, WH04 eliminated the meridional degree of freedom. Omission of this degree of freedom allows the gravest pair of RMM PCs to explain more of the total variance in the resulting reduced basis, but the approach sacrifices detail associated with seasonal and event-to-event variations in the structure of the MJO and its meridional propagation. Concentration of most of the variance associated with the MJO into a single pair of eigenmodes allows the RMM PCs to be plotted in a simple two dimensional phase diagram. The phase concept allows users to sort each day when the signal is deemed present into one of several zonal phases. Identification of the dates characterized by a given phase facilitates the calculation of composite patterns associated with that phase.

In contrast, retention of the meridional and consideration of temporal degrees of freedom as expressed by retention of several leading eigenmodes of RS09 allows their EEOF projection algorithm to more completely specify the temporal and spatial variations of observed convectively coupled modes. Variance associated with the MJO is spread over a large number of the RS09 EEOF PCs. Thus many different combinations of PCs in the EEOF system can explain a given phase of the dual-PC RMM system, each with different characteristics of meridional structure or phase speed. The EEOF PCs thus define a basis much larger than that generated by dual-mode approaches. Spreading some of the variance in the leading eigenmodes across a broader basis implies that a two dimensional phase diagram based on the leading two PCs in the EEOF system does not

include sufficient variance to be as useful as the RMM PCs. Thus the choice between a multi-mode or a two-mode approach to indexing the MJO depends on the ultimate objective of the user.

Development of time indices to represent the MJO and convectively coupled waves allows for empirical prediction of the associated modes. Statistical models can be applied to diagnose and predict the spatial patterns associated with PCs in either the EEOF or the dual PC approaches. Jiang et al. (2008) applied multivariate linear lag regression to predict the RMM PCs and the associated seasonally varying spatial patterns. They incorporated the influence of seasonal variations in the MJO by generating regression coefficients based on OLR and wind data during each calendar month, and then they substituted the observed values of the RMM PCs into the resulting models to obtain the predicted OLR anomalies. They calculated correlations between the predicted and observed values of local OLR anomalies and found that coefficients dropped to below 0.5 after 15 days and 0.2 after 30 days. Their results demonstrated predictions that were better than most numerical weather prediction models prior to 2008 (most of which showed skill extending only to 5-7 days, e.g., Waliser 2005; Seo et al. 2005). They further demonstrated that multiple linear regression produces better forecasts for OLR anomalies associated with the MJO than a sampling of other empirical techniques. Since Jiang et al. (2008), some authors have shown skill in numerical model forecasts of MJO signals that exceeds that of the Jiang et al. benchmark (Bechtold et al., 2008; Seo et al. 2009).

This paper develops an improvement of the Roundy and Schreck (2009) EEOF technique that allows for better spatial and temporal resolution along with better-behaved

PCs. Then it describes a linear regression algorithm similar to that of Jiang et al. (2008) for generating forecasts of modified EEOF PCs in the spectral band of the MJO.

2. Data and Methods

2.1 EEOF Approach

Interpolated OLR data, (Liebmann and Smith, 1996) were obtained from the NOAA Earth System Research Laboratory (ESRL) website. OLR is a relatively good proxy for moist deep convection in the tropics. Although other types of data have proven more effective in recent years in tracking and diagnosing structures of convectively coupled modes, OLR data are available over a longer climatology, making them better suited for analysis of intraseasonal to interannual patterns. These interpolated data are used from June 1974 through 31 December 2003, except for a period of missing data 17 March through 31 December 1978. NOAA uninterpolated OLR data are interpolated linearly in space and utilized for the remainder of the analysis from 2004 through the present. The local mean and seasonal cycle (including its first 3 harmonics) estimated from the period 1974-2006 are subtracted to generate anomalies. A modified version of the algorithm of RS09 is applied to generate EEOF patterns based on the filtered OLR anomalies, and the unfiltered anomalies are applied to generate the associated PCs and to reconstruct filtered data as follows:

1. Filter the interpolated OLR data for a selected band in the zonal wave number frequency domain (filter bands are shown in Fig. 1, except that 100-day low pass filtered anomalies were also analyzed). These bands are broader than those of Wheeler and Kiladis (1999) and Roundy and Frank (2004) in order to include more of the total variance, since the main objective is prediction of sensible weather associated with

equatorial waves, intraseasonal oscillations, and other coherent signals in the same general vicinity of the spectrum.

2. Construct a matrix \mathbf{X} whose columns are time series of the filtered OLR data on the full 2.5° grid from 30°S to 30°N (The original Roundy and Schreck (2009) algorithm applied a reduced grid).

3. Find the leading EOFs \mathbf{E} of the matrix \mathbf{X} (these are the eigenvectors of the matrix $\mathbf{X}^T\mathbf{X}$ corresponding to the largest eigenvalues). Find the corresponding principal component time series (PCs, \mathbf{U}), following

$$\mathbf{U} = \mathbf{X}\mathbf{E}. \quad (1)$$

4. Construct a matrix \mathbf{X}_{pc} from this set of PCs \mathbf{U} , and extend the matrix by including the same PCs at time lags from the original, i.e.,

$$\mathbf{X}_{\text{pc}}(t) = [\mathbf{U}(t), \mathbf{U}(t-1), \mathbf{U}(t-2), \dots, \mathbf{U}(t-n)] \quad (2)$$

5. Calculate the leading EOFs \mathbf{E}_{pc} of \mathbf{X}_{pc} . Steps 1-5 complete the calculation of the extended EOF patterns of the filtered data.

6. Construct a matrix $\mathbf{X}_{\text{unfil}}$ identical to \mathbf{X} except using unfiltered OLR anomaly data, with data reconstructed for lower-frequency modes (if any) calculated first and subtracted. Data included in the 100-day low pass projections are smoothed with a 10-day running mean—otherwise no smoothing is applied.

7. Find

$$\mathbf{U}_{\text{unfil}} = \mathbf{X}_{\text{unfil}}\mathbf{E}. \quad (3)$$

8. Construct the matrix

$$\mathbf{X}_{\text{pcunfil}}(t) = [\mathbf{U}_{\text{unfil}}(t), \mathbf{U}_{\text{unfil}}(t-1), \mathbf{U}_{\text{unfil}}(t-2), \dots, \mathbf{U}_{\text{unfil}}(t-n)] \quad (4)$$

and

9. find

$$\mathbf{U}_{pcunfil} = \mathbf{X}_{pcunfil} \mathbf{E}_{pc}. \quad (5)$$

These reconstructed PCs $\mathbf{U}_{pcunfil}$ are similar to \mathbf{U}_{pc} .

10. An approximation of the filtered data \mathbf{X}_{recon} , applicable in real time, but with reduced noise, is reconstructed by following

$$\mathbf{U}_{recon} = \mathbf{U}_{pcunfil} \mathbf{E}_{pc}, \quad (6)$$

and,

$$\mathbf{X}_{recon} = \mathbf{U}_{recon} \mathbf{E}. \quad (7)$$

The EOF patterns \mathbf{E} and \mathbf{E}_{pc} , form templates of the temporal-spatial patterns of the filtered data onto which the algorithm projects unfiltered data to generate time indices of the modes in each band ($\mathbf{U}_{pcunfil}$) as well as to reconstruct the filtered data in real time without requiring application of the Fourier transform at the end of the data set. Time extension EOF analysis of the original PCs then distinguishes modes that propagate in different directions or with different phase speeds. EOF analysis performed only in space cannot distinguish between eastward and westward propagation if both eastward and westward-moving disturbances have similar spatial structures. Exclusion of higher EOFs reduces noise. Except for the 100-day low pass band, this approach eliminates the need for smoothing the unfiltered data that was necessary following Roundy and Schreck (2009). The algorithm also dramatically reduces the computer memory required while simultaneously increasing the temporal and spatial resolution of the structure of the EEOF patterns identified in each band.

The number of time lags n applied varies across the set of filter bands to save memory. Lower frequency modes include longer histories (a larger n) to better resolve

the modes. The 2-10 day westward band is assigned 12 days, the ER and MJO bands get 100 days, the Kelvin band gets 35 days, and the 100-day low pass band is assigned 1000 days. With the exception of the low pass band, these values match or exceed the longest period included in the band to allow signals in the EOF patterns to resolve at least one cycle. Results are not sensitive to small changes in n . All time lags are applied in increments of 1 day as indicated in (2) and (4), except for the low pass band, which uses 10-day time stepping.

Objective determination of the appropriate number of EOFs to retain is difficult and might not improve upon subjective approaches. Several tens of EOFs are required to explain most of the variance in each band. These are taken from a basis of greater than 14,000 possible, so that those retained are a tiny fraction of the total. Traditionally, retaining several tens of EOFs is considered bad practice. However, retaining just a handful is insufficient to resolve the spatial structures of observed convective disturbances. To simplify the analysis of results, the number of EOFs retained was arbitrarily set at 75% of the total variance. This setting provides a reasonable representation of the original filtered data, especially across the warm pool zones of the tropical oceans (and the Pacific Ocean cold tongue for the low pass band) while still discarding a substantial amount of noise. Dramatic reduction in the amount of variance retained reduces the effective resolution of the structure and evolution of observed large-scale convective systems. For the remainder of this analysis, I retain enough eigenmodes to explain 75% of the variance in each band, and I refer to the result as the “signal” and the remaining 25% of the variance as “noise”. The number of EOFs retained is 82, 39, 262, 122, and 38 for the ER, MJO, 2-10 day westward, Kelvin, and 100-day low pass

bands, respectively. In general, bands with signals characterized by larger wave numbers and higher frequencies require greater numbers of EOFs to resolve the observed signals, consistent with the larger number of degrees of freedom expected in such bands. Although some aspects of the approach are arbitrarily defined, the analysis below shows that the resulting signals represent well patterns in unprocessed OLR anomalies, and that prediction of the extracted signals associated with the MJO is skillful in independent data past one month in some regions of the tropics.

2.2 Forecasting by Multiple Linear Regression

This section discusses an approach for predicting projected MJO band OLR anomalies discussed in Section 2.1, step 10. The first step applies the matrix of principal components for the MJO band $\mathbf{U}_{pcunfil}$ in a multiple linear regression model to predict itself at a time lag τ :

$$\mathbf{U}_{pcunfil}(t + \tau) = \mathbf{U}_{pcunfil}(t)A_{\tau}, \quad (8)$$

where A_{τ} is a vector of regression coefficients. Equation (8) is solved for A_{τ} including only days of the year 45 days before to 45 days after the day of the year on which the intended forecast is made. Focus on one time of the year allows the algorithm to better forecast the evolution of patterns in the selected wave-number frequency band that evolve differently during different seasons (similar to Jiang et al. 2008, who calculated regression coefficients in their approach according to calendar month). The training period includes all available data from 27 March 1975 through 8 April 2007. That period represents 11,700 days, or 2,925 days for each regression calculation (after accounting for the requirement for day of year).

Steps 9 and 10 in Section 2.1 reconstruct the filtered OLR anomalies corresponding to the forecast. Skill is assessed by cross-validation to generate a blind hind cast dataset. Hind casts are made for each day of each year, based on training the regression models on data from all other years. This validation approach is complicated because the original filtered data and the EEOF patterns themselves contain temporal information that might artificially enhance the apparent skill of the hind casts. To address this problem, new EEOF patterns and PCs were calculated for each year of the dataset. The calculations for a given year are made by removing the data for that year from the calculation of the EEOF patterns that are then applied for predicting signals that year. This approach results in fragmentation of the dataset that might reduce the skill of prediction, but tests reveal that the basis of structures in the basis of EEOFs change little from year to year in the MJO band.

Roundy and Schreck (2009) showed that nearly every EEOF pattern in their analysis appears as a pair of eigenmodes that are in temporal and spatial quadrature, and members of each pair explain nearly the same amount of variance. This conclusion also applies to the revised algorithm considered here. Since removal of a single year can redistribute small amounts of variance between the eigenmodes, the order of eigenvectors is not maintained in the calculations for each year. Thus this cross validation approach confounds the testing of the skill of prediction of individual PCs. However, the projected OLR anomalies associated with combinations of many PCs do not appear to be significantly affected. This project therefore determines skill of the forecasts through analysis of the predicted projected OLR fields. Since the number of EOFs retained is arbitrarily set to retain 75% of the variance and the testing dataset is not applied to select

predictors, artificial skill due to predictor screening (DelSole and Shukla 2009) does not contribute to the results. Artificial skill associated with a large number of predictors also does not confound this analysis because the results are tested with independent data.

3. Results

3.1. Projected OLR Signals

Sections 3.1-3.3 analyze the projected signals and their similarity to the filtered data. Figure 2 shows an example of unfiltered OLR anomalies (shading) and projected OLR signals (contours), averaged from 2.5°N to 12.5°N for January through July 1997. Blue shades suggest active convection. Solid contours represent negative projected OLR anomalies, with dashed positive. Black represents the 100-day low pass projections, red, the MJO band, green, the ER wave band, magenta, the Kelvin band, and cyan the 2-10 day westward band. During the first half of the period, low frequency convection is suppressed over the western Indian Ocean basin and enhanced over the western Pacific and Atlantic Oceans. A transition to a pattern consistent with El Niño occurred across April, with enhanced low frequency convection indicated over the Pacific Ocean and suppressed low frequency convective signals developing over the Maritime Continent. Red contours indicate alternating periods of enhanced and suppressed convection in the MJO band, with the strongest signals across the Indian Ocean and Maritime continent regions. Some MJO band signals also occur over the Atlantic Ocean and Africa. Each MJO event evolves differently, but each anomaly represented by red contours appears to be associated with substantial signals in unfiltered OLR anomalies. Green contours representing equatorial Rossby wave yield similar conclusions. Fast signals associated with Kelvin waves circle the globe as indicated by the magenta contours. Extratropical

wave signals can also project substantially onto signals in the Kelvin band, but extratropical waves and Kelvin waves are not independent (Straub and Kiladis 2003). Easterly waves, mixed Rossby gravity waves, and some tropical cyclone signals appear together in the blue contours that represent the 2-10 day westward band. Overall, the contoured signals appear to represent well the target disturbances in the unfiltered data.

3.2. Spectrum Analysis

A zonal wave number frequency spectrum analysis is applied to the projected OLR anomalies in the ER, MJO, 2-10 day westward, and Kelvin bands. Projected data obtained from Section 2.1 step 10 are sorted into longitude time arrays at each latitude from 15°S to 15°N. Each array is then broken into a series of 360-day segments overlapping each other by 180 days. The ends of each segment are then tapered in time by multiplication by cosine bells to reduce spectral leakage. The overlap recovers data lost to the tapering. This general approach is similar to that followed by Wheeler and Kiladis (1999), but with longer time windows. The result is then normalized for plotting by dividing by the product of the number of days in each window and the number of zonal grid points. No attempt to remove a background is applied because noise is already reduced by exclusion of the EEOFs associated with the bottom 25% of the variance in each band.

Figure 3 shows the resulting spectra for the ER, MJO, 2-10 day westward, and Kelvin bands in panels a-d, respectively. Although slight overlap in power occurs between the ER and 2-10 day westward band and between the MJO and Kelvin bands, this analysis demonstrates that the projection algorithm separates the signals well. Any overlap is not generated by redundant signals, because the ER projections and the MJO projections are

removed from the data before generating the 2-10 day westward and Kelvin projections. Power in each band tends to concentrate in the lower frequencies of the band. Previous works have associated this concentration with a red noise background. However, these results suggest that coherent signals in the bands are also red.

The spectral peak of projected signals in the Kelvin band (Fig. 3 d) spreads across a broad range of wave numbers at each frequency. Previous works have suggested that the spectral peak associated with Kelvin waves is narrow in wave number relative to the peak associated with the MJO. Power extending from the core of the Kelvin peak to higher wave number has generally be assumed to be part of the red background. However, Fig. 3d suggests that a portion of that power is associated with coherent signals. Thus the peak associated with Kelvin waves may in fact be broader than previously thought, and less inconsistent with the MJO peak. This analysis also suggests that substantial coherent signals occur in the “spectral gap” between the MJO and Kelvin peaks, where power has previously been assumed to be part of the background. These results raise questions about the previous assumptions that power above a smoothed background is a necessary condition for the presence of coherent signals. A portion of that smoothed background spectrum is generated by seasonal and geographical changes in the propagation characteristics of the different modes. For example, Roundy (2008) showed that Kelvin waves tend to propagate more quickly across the western hemisphere than the eastern, and they tend to slow down over the eastern hemisphere as they propagate through the local active convective phase of the MJO and speed up through the suppressed phase. As different events evolve differently, their contributions to the spectrum appear at different wave numbers and frequencies, leading to a smoother spectrum with power distributed

more evenly. Such changes in wave patterns would affect sensible weather conditions and ought not be discarded. Applying broad filter bands to process OLR data would retain such variations, while narrow filters like those of Wheeler and Kiladis (1999) and Roundy and Frank (2004) would exclude them. Those previous works benefited from the narrow filter bands because they helped demonstrate the clear association between some observed signals and the dispersion characteristics of theoretical waves. In contrast, the core objective of the present analysis is to assess and predict coherent signals associated with more of the variance in broader neighborhoods of the full spectrum.

3.3 Comparison of Projected and Wavenumber-Frequency Filtered OLR

Figure 4 shows the temporal Pearson correlation coefficients of the filtered and the projected OLR anomalies at each grid point across the tropics. Correlation coefficients exceed 0.9 in regions of maximum OLR variance in the same wave number frequency bands (Roundy and Frank, 2004). Correlations as low as 0.3 occur in regions of climatologically low OLR variance, such as the region of the subtropical ridge west of South America. Figure 5 shows the corresponding fraction of the local variance in the filtered OLR explained by the projected OLR. As with the correlation analysis, the fraction of the local variance of filtered OLR explained by the projected OLR is greatest in regions where OLR in the target wave number frequency band varies most. The projected data explain less than 10% of the local variance in some geographical regions where OLR anomalies in the target bands tend to be small. These results suggest that most of the noise discarded in the projection process originates from regions with climatologically low variance. Some of this “noise” might be associated with real weather events. However, consistent with the Roundy and Schreck (2009), these results

suggest that the EEOF projection algorithm reduces some of the noise associated with filtering in the wave number frequency domain, such as ringing of signals from regions of high activity into regions of climatologically low activity. The projected signals thus might provide a better representation of actual coherent patterns in unfiltered data than filtered data.

3.4 Example Hindcasts

Figure 6 shows projected OLR signals in the MJO band (shading), with contours representing blind hind casts of the same anomalies for lead times of 7, 14, 21, and 28 days (shown in panels a-d, respectively). Blue shades suggest active convection, but forecast active convective signals are contoured in red to enhance the contrast. The 7-day forecast indicates high amplitude signals nearly identical to the verification. Although the forecast signal degrades with higher lead times, some patterns remain well represented even at 28 days. On the other hand, some hind cast anomalies at 21 and 28 day lead times suggest outcomes opposite the verification, especially during April through June 1988.

3.5 Assessment of Skill

Cross-validated hind casts were generated for projected OLR anomalies at daily lead times up to 30 days for 1974 through 2006. The standard deviations, RMS errors, and correlations were calculated for the hind casts at each lead-time in both space and time across 30°S to 30°N and select geographical regions. Verification data are the OLR projections onto the EEOFs of the MJO band. A similar assessment of the OLR hind casts following Jiang et al. (2008) was prepared for comparison, based on the same EEOF verification benchmark, for consistency.

The Taylor diagram (Taylor 2001) provides a convenient assessment of the skill of a forecast. This diagram has become popular among climate scientists to demonstrate the quality of climate model output relative to observed data. The Taylor diagram provides information about the correlation between the forecast and verification data, as well as about the amplitude of the forecast signal and its root mean square error. Figure 7 shows the Taylor diagram for 30°S to 30°N hind cast MJO band OLR for the EEOF forecast (red) and the Jiang et al. (2008) approach (blue). The standard deviation of the observed signal is 6 Wm^{-2} , and a forecast in the lower right corner of the diagram would have perfect correlation and zero RMS error. The EEOF forecast data always explain more of the variance in the projected OLR anomalies than the Jiang et al. hind casts. This comparison might not seem fair, but it is important to point out that the EEOF projected data set explains 50% more of the variance in unfiltered OLR anomalies than that obtained following Jiang et al. (2008). This larger amount of variance explains most of the improvement of the EEOF approach over the Jiang et al. approach. The EEOF forecasts lose correlation with lead-time more quickly than they lose amplitude. Nevertheless, correlations remain significant to the 25-day lead. RMS error never exceeds the standard deviation of the verification data at any lead. The Jiang et al. approach shows slightly lower RMS error after the 24-day lead than the EEOF forecast. Consistent with the above interpretation of Fig. 7, the EEOF forecast approach shows some skill across the global tropics approaching 25-day lead times.

Figure 8 shows a similar Taylor diagram for the EEOF approach including only India (15 to 25°N and 65 to 85°E). Correlations and error statistics include only June 1 through August 31, to assess the skill of the EEOF algorithm in predicting the Indian southwest

monsoon. The 30-day forecast has a correlation of roughly 0.5 and it retains roughly half the standard deviation of the verification data. These results suggest that skill in predicting MJO band signals in the Indian summer monsoon extends to at least 30 days. All correlations plotted in Fig. 8 are significant at above the 99% level.

These results suggest that correlation and error evolve differently for different geographical locations and times of the year. The skill of the forecast might also vary with the amplitude of the predicted signal. Figure 9a shows the correlation skill assessed at each latitude for the region between 40°E and 90°E including only times when the forecast signal at that lead-time exceeds +/- 1 standard deviation in the cross validated forecast data. Correlation drops off more quickly near the equator than at high latitude, with correlations remaining statistically significant poleward of 15°N or S beyond 30-day lead times. Figure 9b shows the corresponding skill score (SS) relative to climatology (the zero anomaly):

$$SS = -\frac{RMSE - RMSE_{clim}}{RMSE_{clim}}. \quad (9)$$

Skill declines with time to 0 on the equator by 26 days, but remains positive farther off the equator through 30 days. Skill drops to zero within 10 days in a similar analysis of SS for times when the forecast suggests amplitudes less than 1 SD (not shown). Thus high amplitude forecast signals can be taken with confidence, but forecasts of low amplitude signals might not be as useful.

Figure 9 suggests that skill decreases smoothly with increasing lead-time from the equator toward the poles. This pole ward migration of skill might be associated with pole ward-moving signals linked with the MJO during some times of the year such as the northern summer over Asia and the western north Pacific Ocean or the southern summer

over the southwest Pacific Ocean. Thus the present state of MJO band convection yields more information about convection at high latitudes of the tropics past week 3 than it yields about future convection near the equator at the same time. This result suggests that the algorithm might be less effective at forecasting new events near the equator than forecasting the continued pole ward drift of existing events.

4. Conclusions

This work describes a new algorithm for diagnosing signals from satellite OLR data associated with convectively coupled waves, intraseasonal oscillations, and climate variations. Signals in selected broad bands of the zonal wave number frequency domain are associated with temporal spatial eigenvectors and principal components that can be applied to dissect the patterns of convection in real time and to predict their temporal and spatial evolution. The principal components serve as time indices of signals in the selected bands. The original filtered data are reconstructed from unfiltered data by projecting unfiltered data onto the time-space eigenmodes of the filtered data. As presently applied in real time at

<http://www.atmos.albany.edu/facstaff/roundy/waves/hovsdet/>

and

<http://www.atmos.albany.edu/facstaff/roundy/waves/plotsmovrec523.html>,

the resulting reconstructed filtered fields explain roughly 75% of the variance in the corresponding filtered data and the results represent well the signals of the filtered data in the geographical regions where the filtered signals vary most. The PCs in a given band are easily predicted by multiple linear regression based on regression coefficients calculated from data at the same time of the year as the forecast.

Although the approach is applied at the same website to predict signals in all five wave number frequency bands analyzed here, for simplicity, this paper assesses only the skill of the forecasts for signals in the MJO band because these signals are of greatest interest. The forecast MJO band signals exhibit some skill to about 25 days when that skill is assessed across the global tropics and more substantial skill past 30 days for strong predicted signals at high latitudes of the tropics (such as India and over the northwest tropical Pacific Ocean). Analysis of the hind casts of signals in the other bands will be included in a future manuscript.

Acknowledgements: The NOAA Earth System Research Laboratory graciously provided OLR data for this analysis. Funding was provided by the National Science Foundation, grant 0850642.

Figure Captions

Figure 1 Boxes represent bands of the zonal wave number frequency domain applied to filter OLR anomalies for EEOF pattern extraction. The bands are plotted on a wave number frequency power spectrum of OLR anomalies normalized by dividing by a smoothed background spectrum, calculated as by Roundy (2008). Thin gray curves represent dispersion solutions for waves on the equatorial beta plane (see Kiladis et al. 2009) for the indicated equivalent depths (h).

Figure 2: OLR anomalies (shading, Wm^{-1}) averaged from 2.5°N to 12.5°N for January through July 1997. Contours represent EEOF projected OLR anomalies, with solid contours negative and dashed positive. Black represents 100-day low pass signals, and red, green, magenta, and cyan represent the MJO, ER, Kelvin, and 2-10 day westward bands, respectively. The contour interval is 7.5 Wm^{-2} , and the zero contour is omitted.

Figure 3: Power spectra of projected OLR anomalies in the indicated bands.

Figure 4: Local correlation between filtered and projected OLR anomalies for the bands indicated in the panel titles.

Figure 5: Fraction of the local variance in OLR anomalies filtered for the indicated bands of the wave number frequency domain explained by the corresponding EEOF projected OLR anomalies.

Figure 6: Verification EEOF projected OLR anomalies in the MJO band (shading, with active convection suggested in blue), and the corresponding cross-validated predicted signal at lead times of 7, 14, 21, and 28 days in panels a-d

(respectively). Red contours indicated negative anomalies and blue contours positive. The contour interval is 5 Wm^{-2} , and the zero contour is omitted.

Figure 7: Taylor diagram representing the skill of the hind cast OLR anomalies in the MJO band. Red indicates the result for the EEOF approach, and blue represents the result from the comparable forecast of Jiang et al. (2008). Numbers of the same color near the plotted points indicate the lead time of the forecast represented at that number. A perfect forecast would have an RMSD of 0, correlation of 1, and a standard deviation of 6.

Figure 8: Taylor diagram as in Fig. 7, but correlations, standard deviations, and RMS errors are calculated only for the region 15°N to 25°N and 65°E to 85°E during June, July, and August to demonstrate skill of the EEOF technique in predicting MJO band OLR anomalies during the Indian southwest monsoon.

Figure 9 a. Pattern correlation between EEOF forecast and verification OLR anomalies in the MJO band for the region 40°E to 90°E as a function of latitude and lead-time, including only times when the forecast indicates signals in excess of ± 1 SD at the same lead times in the cross-validated hind cast data set. b. The corresponding skill score (SS) is defined in equation 9. Minimum skill is -0.015. Contours are plotted every 0.05 for both panels a and b.

Works Cited

- Bechtold P, Kohler M, Jung T, Doblas-Reyes F, Leutbecher M, Rodwell MJ, Vitart F, Balsamo G. 2008. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Q. J. Roy. Meteorol. Soc.* 134:1337-1351. DOI:10.1002/qj.289.
- Branstator GW. 2002. Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation. *J. Climate*, 15:1893-1910.
- DelSole T, Shukla J. 2009. Artificial skill due to predictor screening. *J. climate*. 22:331-345. doi: 10.1175/2008JCLI2414.1.
- Ferranti L, Palmer TN, Molteni F, and Klinker K. 1990. Tropical-extratropical interaction associated with the 30-60 day oscillation and its impact on medium and extended range prediction. *J. Atmos. Sci.* 47:2177-2199.
- Gottschalk J, Wheeler M, Weickmann K, Vitart F, Savage N, Hendon HH, Lin H, Flatau M, Waliser D, Sperber K, Higgins W, and Vintzileos A. 2010. Establishing and assessing operational model MJO forecasts: A project of the CLIVAR Madden-Julian Oscillation working group. Submitted to *Bull. Am. Meteor. Soc.*
- Hendon HH, Liebmann B, Newman M, Glick JD, and Schemm JE. 2000. Medium-range forecast errors associated with active episodes of the Madden-Julian oscillation. *Mon. Wea. Rev.* 128:69-86.
- Higgins RW, Schemm J-KE, Shi W, Leetmaa A. 2000. Extreme precipitation events in the western United States related to tropical forcing. *J. Climate*. 14:403-417.
- Jiang X, Waliser D, Wheeler MC, Lee M-I, Schubert SD. 2008. Assessing the skill of an all-season statistical forecast model for the Madden-Julian oscillation. *Mon. Wea. Rev.*, 136:1940-1956.
- Jones C, Schemm J-KE. 2000. The influence of intraseasonal variations on medium-range weather forecasts over South America. *Mon. Wea. Rev.*, 128:486-494.
- Jones C, Waliser DE, Lau KM, Stern W. 2004a. Global occurrences of extreme precipitation and the Madden-Julian Oscillation: Observations and predictability. *J. Climate*, 17:4575-4589.

- Jones, C, Waliser DE, Lau KM, Stern W. 2004b. The Madden-Julian oscillation and its impact on Northern Hemisphere weather predictability. *Mon. Wea. Rev.*, 132:1462-1471.
- Kiladis, GN, Straub KH, Haertel PT. 2005. Zonal and vertical structure of the Madden-Julian Oscillation. *J. Atmos. Sci.*, 62:2790-2809.
- Kiladis GN, Wheeler MC, Haertel PT, Straub KH, Roundy PE. 2009. Convectively coupled equatorial waves. *Reviews of Geophysics*, 47:RG2003, doi:10.1029/2008RG000266.
- Liebmann, B, Smith CA. 1996. Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull. Amer. Meteor. Soc.* 77:1275-1277.
- Madden RA, Julian PR. 1994. Observations of the 40-50-day tropical oscillation-A review. *Mon. Wea. Rev.* 122:814-837.
- Mo, KC, Higgins RW. 1998. Tropical influences on California precipitation. *J. Climate.* 11:412-430.
- Mo, KC. 2000. Intraseasonal modulation of summer precipitation over North America. *Mon. Wea. Rev.* 128:1490-1505.
- Nogues-Paegle J, Byerle LA, Mo KC. 2000. Intraseasonal modulation of South American summer precipitation. *Mon. Wea. Rev.* 128:837-850.
- Roundy PE, WM Frank. 2004. A climatology of waves in the equatorial region. *J. Atmos. Sci.* 61:2105-2132.
- Roundy PE, MacRitchie K, Asuma J, Melino T. 2010. Modulation of the Global Atmospheric Circulation by Combined Activity in the Madden Julian Oscillation and the El Niño/Southern Oscillation During Boreal Winter. *J. Climate.* 23:4045-4059.
- Roundy PE, Schreck CJ. 2009. A combined wave-number-frequency and time-extended EOF approach for tracking the progress of modes of large-scale organized tropical convection. *Q. J. Roy. Meteorol. Soc.* 135:161-173.

- Seo K-H, Schemm J-K, Jones C. 2005. Forecast skill of the tropical intraseasonal oscillation in the NCEP GFS dynamical extended range forecasts, *Climate Dyn.* 25(2-3):265-284.
- Seo K-H, Wang W, Gottschalck J, Zhang Q, Schemm J-KE, Higgins WR, Kumar A. 2009. Evaluation of MJO forecast skill from several statistical and dynamical forecast models. *J. Climate.* 22:2372-2388. doi: 10.1175/2008JCLI2421.1.
- Straub KH, Kiladis GN. 2003. Extratropical forcing of convectively coupled Kelvin waves during austral winter. *J. Atmos. Sci.* 60:526-543.
- Taylor K. 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* 106:D7.
- Vitart F, Molteni F. 2009. Simulation of the MJO and its teleconnections in an ensemble of 46-day EPS hindcasts. *ECMWF Technical Memo*, **597**, 60 pp, available at <http://www.ecmwf.int/publications/library/do/references/list/14>.
- Waliser D. 2005. Predictability and forecasting. *Intraseasonal Variability of the Atmosphere-Ocean Climate System*. Lau WKM, Waliser DE, Eds., Springer-Verlag, 389-423.
- Wallace JM, Gutzler DS. 1981. Teleconnections in the geopotential height field during Northern Hemisphere winter. *Mon. Wea. Rev.*, 109:784-812.
- Weickmann KM, Berry E. 2007. A synoptic-dynamic model of subseasonal atmospheric variability. *Mon. Wea. Rev.* 135:449-474.
- Weickmann KM, Kiladis GN, Sardeshmukh PD. 1997. The dynamics of intraseasonal atmospheric angular momentum oscillations. *J. Atmos. Sci.* 54:1445-1461.
- Wheeler M, Hendon HH. 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.* 132:1917-1932.
- Zhang C. 2005. The Madden-Julian Oscillation. *Rev. Geophys.*, 43:RG2003, doi:10.1029/2004RG000158.

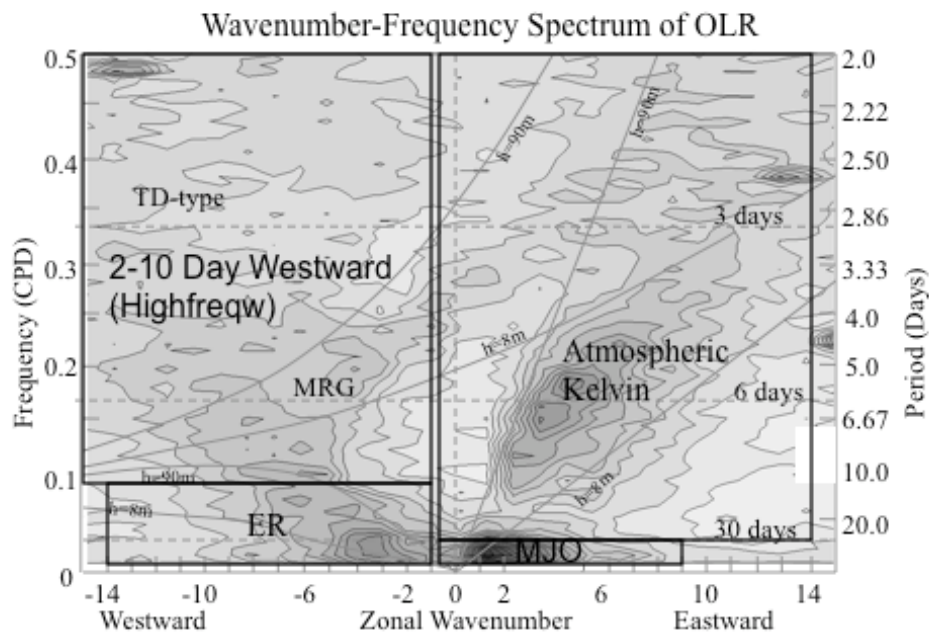


Figure 1 Boxes represent bands of the zonal wave number frequency domain applied to filter OLR anomalies for EEOF pattern extraction. The bands are plotted on a wave number frequency power spectrum of OLR anomalies normalized by dividing by a smoothed background spectrum, calculated as by Roundy (2008). Thin gray curves represent dispersion solutions for waves on the equatorial beta plane (see Kiladis et al. 2009) for the indicated equivalent depths (h).

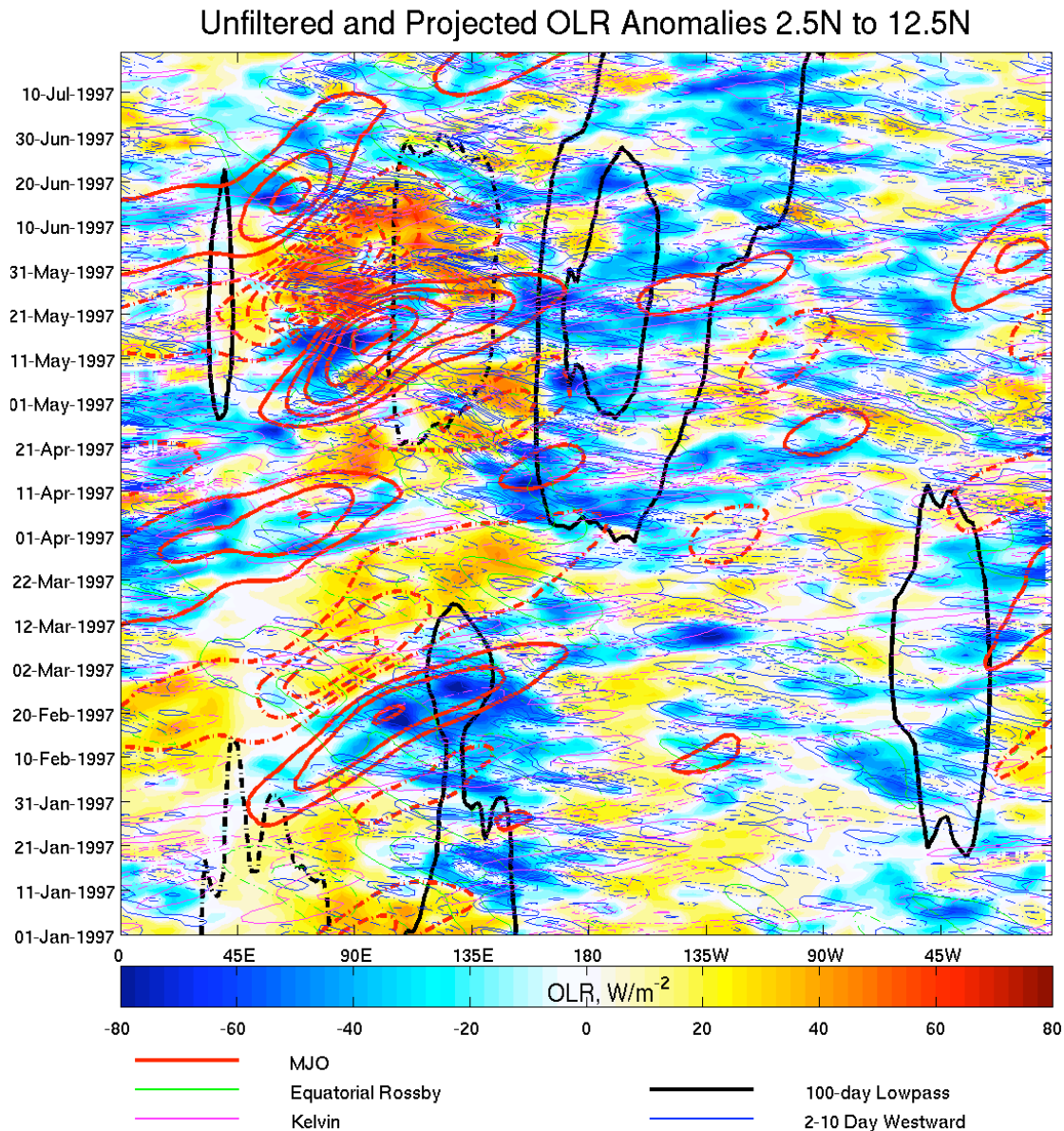


Figure 2: OLR anomalies (shading, Wm^{-1}) averaged from $2.5^{\circ}N$ to $12.5^{\circ}N$ for January through July 1997. Contours represent EEOF projected OLR anomalies, with solid contours negative and dashed positive. Black represents 100-day low pass signals, and red, green, magenta, and cyan represent the MJO, ER, Kelvin, and 2-10 day westward bands, respectively. The contour interval is $7.5 Wm^{-2}$, and the zero contour is omitted.

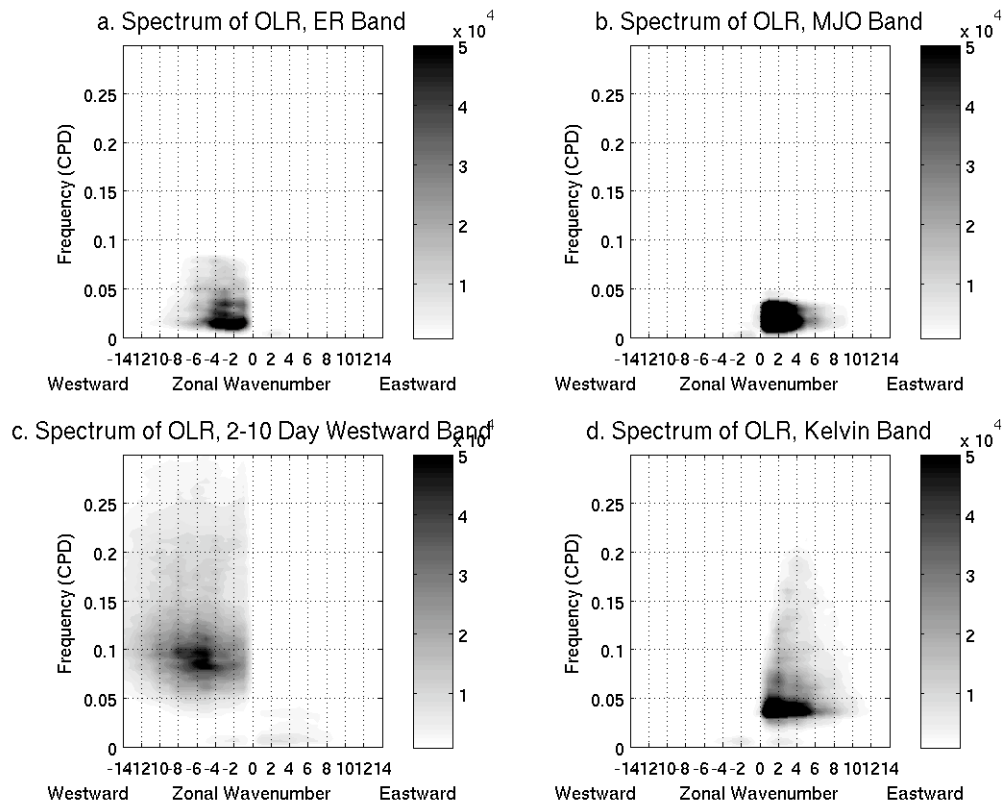
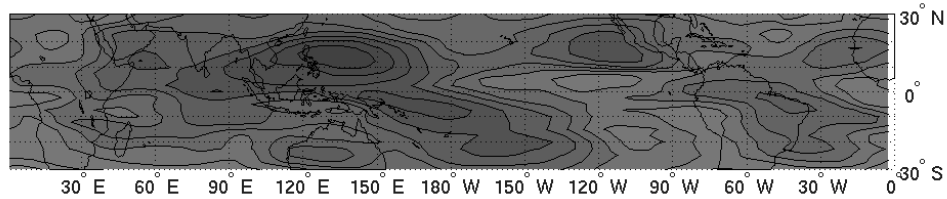
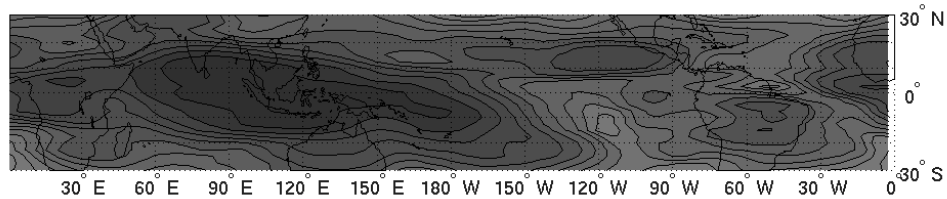


Figure 3: Power spectra of projected OLR anomalies in the indicated bands.

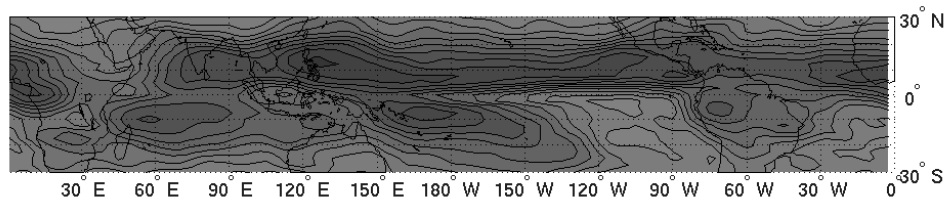
a. Correlation between Projected and Filtered OLR in the ER band



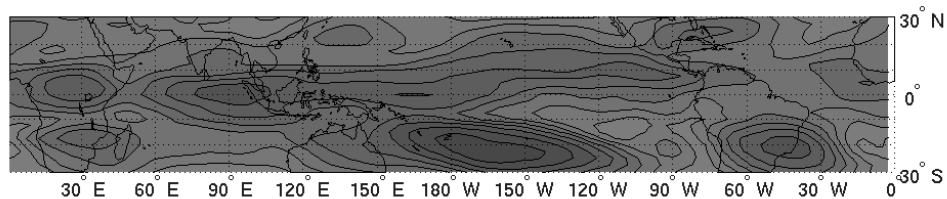
b. Correlation between Projected and Filtered OLR in the MJO band



c. Correlation between Projected and Filtered OLR in the 2-10 Day Westward band



d. Correlation between Projected and Filtered OLR in the Kelvin band



e. Correlation between Projected and Filtered OLR in the lowpass band

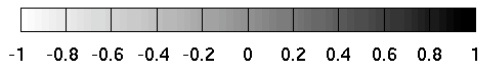
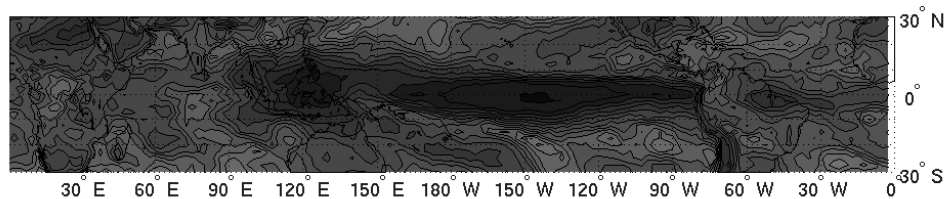


Figure 4: Local correlation between filtered and projected OLR anomalies for the bands indicated in the panel titles.

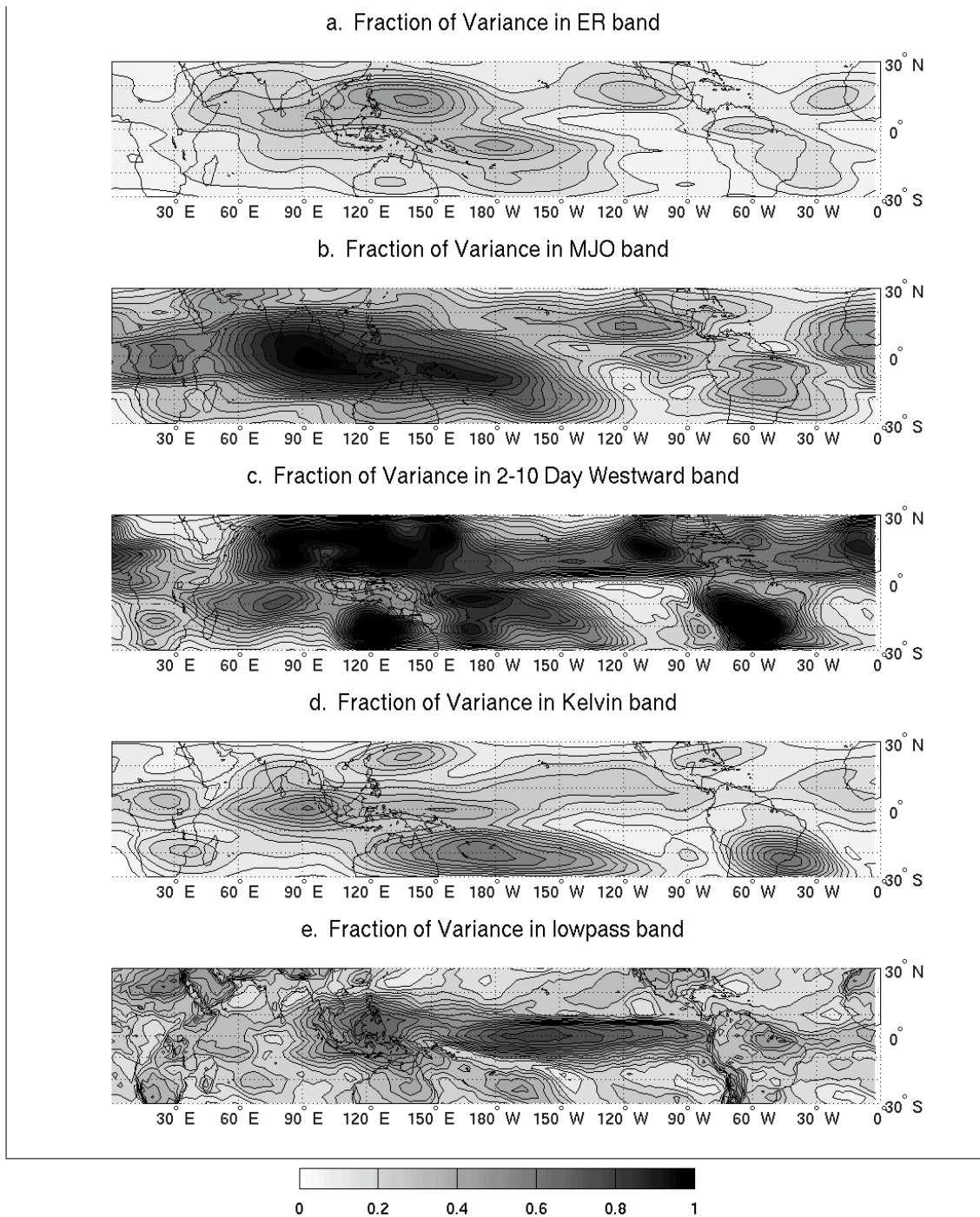


Figure 5: Fraction of the local variance in OLR anomalies filtered for the indicated bands of the wave number frequency domain explained by the corresponding EEOF projected OLR anomalies.

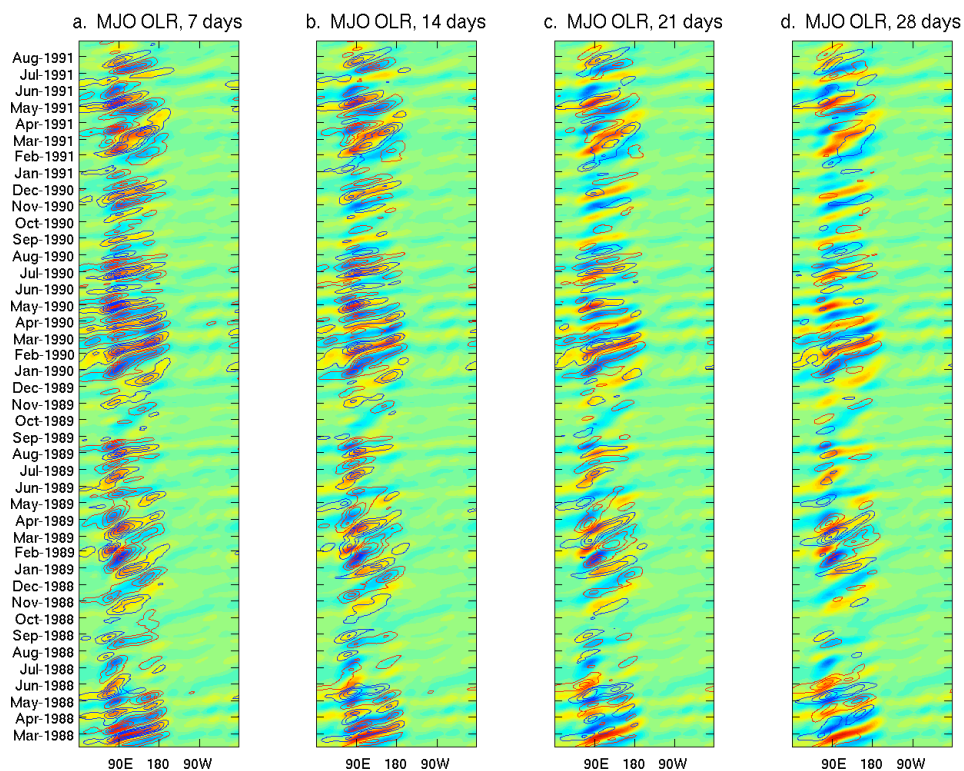


Figure 6: Verification EEOF projected OLR anomalies in the MJO band (shading, with active convection suggested in blue), and the corresponding cross-validated predicted signal at lead times of 7, 14, 21, and 28 days in panels a-d (respectively). Red contours indicated negative anomalies and blue contours positive. The contour interval is 5 Wm^{-2} , and the zero contour is omitted.

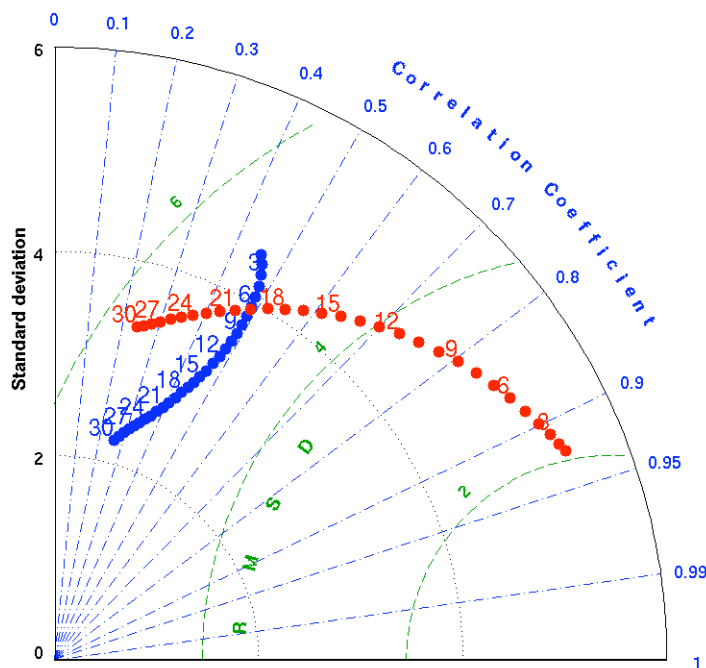


Figure 7: Taylor diagram representing the skill of the hind cast OLR anomalies in the MJO band. Red indicates the result for the EEOF approach, and blue represents the result from the comparable forecast of Jiang et al. (2008). Numbers of the same color near the plotted points indicate the lead time of the forecast represented at that number. A perfect forecast would have an RMSD of 0, correlation of 1, and a standard deviation of 6.

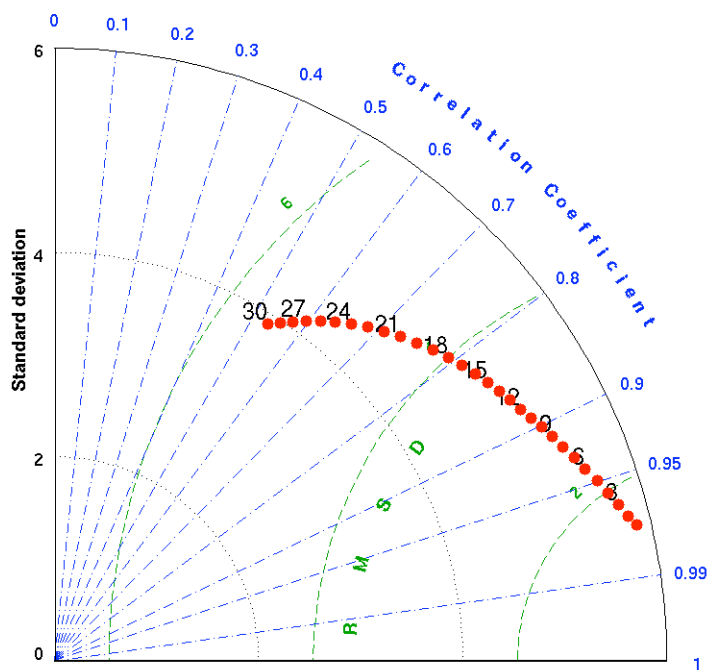


Figure 8: Taylor diagram as in Fig. 7, but correlations, standard deviations, and RMS errors are calculated only for the region 15°N to 25°N and 65°E to 85°E during June, July, and August to demonstrate skill of the EEOF technique in predicting MJO band OLR anomalies during the Indian southwest monsoon.

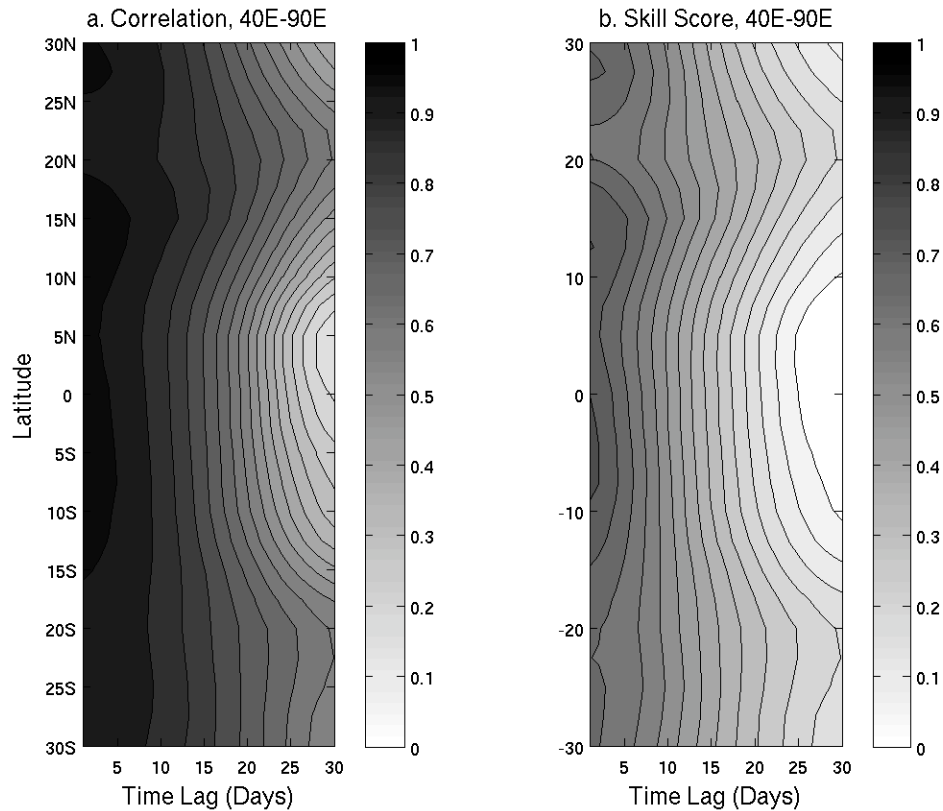


Figure 9 a. Pattern correlation between EEOF forecast and verification OLR anomalies in the MJO band for the region 40°E to 90°E as a function of latitude and lead-time, including only times when the forecast indicates signals in excess of ± 1 SD at the same lead times in the cross-validated hind cast data set. b. The corresponding skill score (SS) is defined in equation 9. Minimum skill is -0.015 . Contours are plotted every 0.05 for both panels a and b.